REMARKS

In the Office Action, the Examiner rejected Claims 1-15, 19-23 and 27-29, which were all of the then pending claims, under 35 U.S.C. 103 as being unpatentable over U.S. Patent 6,405,161 (Goldsmith) in view of U.S. Patent 7,103,536 (Kanno) and further in view of U.S. patent application publication no. 2003/0105638. It is noted that the previous rejection of Claims 1-11, 13-15 and 21-23 under 35 U.S.C. 102 as being fully anticipated by Goldsmith was withdrawn.

Independent Claims 1, 13 and 21 are being amended to better define the subject matters of these claims. In particular, these claims are being amended to describe in more detail the way in which previously discovered knowledge is used to refine the sets of candidate prefixes and suffixes Claim 29 is being amended to describe a feature of an embodiment of the invention.

For the reasons discussed below, Claims 1-15, 19-23 and 27-29 patentably distinguish over the prior art and are allowable. The Examiner is thus asked to reconsider and to withdraw the rejection of Claims 1-15, 19-23 and 27-29 under 35 U.S.C. 103, and to allow these claims.

Generally, Claims 1-15, 19-23 and 27-29 patentably distinguish over the prior art because the prior art does not show or suggest refining sets of potential prefixes and suffixes, using knowledge of previously identified prefixes to further refine a set of candidate suffixes, and using knowledge of previously identified suffixes to further refine a set of candidate prefixes, as described in independent Claims 1, 13 and 21.

In order to best understand this feature and its significance, it may be helpful to review briefly the present invention and the prior art.

Applicants' invention, generally, relates to automatically collecting affixes of a language from one or more documents. Knowledge of affixes is important for analyzing existing words and for producing new words. It is very difficult and time extensive to acquire a complete list of

I:\IBM\105\16751\amend\16751.nm3.doc

affixes of a language by hand, and a number of procedures have been tried to use a computer or some other automated process to identify affixes.

The previous approaches tend to parse words into pieces, either a prefix and a stem, or a stem and a suffix. Also, the prior art approaches tend to limit the length of an affix to reduce the size of the search space.

The prior art systems have a number of disadvantages and limitations. For example, they may fail to discover both prefixes and suffixes at the same time, and they may not be able to discover nested affixes. In addition, because of the limitations on length, the prior art systems may not find many affixes that appear in technical documents. Further, many of the previous approaches fail to find affixes containing non-alphabet characters such as digits and hyphens.

The present invention effectively addresses these prior art limitations. Generally, the present invention provides an unsupervised, knowledge-free procedure for automatically discovering prefixes and suffixes from text. The present invention integrates prefix and suffix discovery in such a way that uses knowledge about prefixes to find suffixes and uses knowledge about suffixes to find prefixes.

More specifically, in one embodiment, the invention provides a computer system for analyzing text in one or more documents, and comprising one or more system interfaces, and an affix process that determines one or more affixes of one or more words in one or more of the documents and provides the affixes to the system interface. This affix determining process comprises obtaining a collection of words, adding the words into a prefix Patricia tree, using said prefix tree to identify a set of candidate prefixes, reversing each of the words, adding the reversed words into a suffix Patricia tree, and using the suffix tree to identify a set of candidate suffixes.

I:\IBM\105\16751\amend\16751.am3.doc

The affix determining process further comprises refining the sets of candidate prefixes and suffixes, using knowledge, previously discovered in the refining step, to identify actual prefixes and suffixes. In particular, knowledge of prefixes previously identified in the refining is used to further refine the set of candidate suffixes, and knowledge of suffixes previously identified in the refining is used to further refine the set of candidate prefixes.

The references of record do not disclose or suggest refining sets of candidate prefixes and affixes as described above.

For instance, Goldsmith discloses an automated, morphological analysis of a natural language for determining prefixes, suffixes and stems. This analysis has three major components: determining the correct morphological split for individual words, establishing accurate categories of stems based on the range of suffixes that they accept, and identifying allomorphs of the same stem.

As the Examiner has recognized, there are a number of important features of the preferred embodiment of this invention that are not shown in or suggested by Goldsmith. For example, Goldsmith does not disclose or render obvious the user of reverse words for identifying suffixes, or the use of Patrician trees to identify candidate sets of prefixes and suffixes. Also, Goldsmith does not disclose or render obvious refining the sets of candidate prefixes and suffixes by using previously identified suffixes and prefixes, respectively.

This feature of Applicants' invention also is not shown in or rendered obvious by either Kanno or Taira.

Kanno describes a process for identifying a suffix. The Examiner cited this reference for the disclosure of reversing a word and adding it to a Patricia tree for suffix matching. There is no teaching, though, in Kanno of the above-described refinement process used in the present invention.

I:\IBM\105\16751\amend\16751.am3.doc

Taira discloses a method and system for generating structured medical reports from natural language reports. In the procedure described in Taira, statistical natural language processing is used to analyze individual words and combinations of words to generate more accurate translations. Paragraph (74) of Taira refers to a recursive algorithm, in which the results of one iteration are sent to the next stage of processing, and based on the results of this stage, may trigger a refinement of a word's classification.

Thus, while Taira discloses the use of the results of one stage to refine the classification of a word, there is now disclosure in this reference of the specific type of refinement used in the present invention – that is, using previously identified <u>prefixes</u> to refine a set of <u>candidate suffixes</u>, or of using previously identified <u>suffixes</u> to refine a set of <u>candidate prefixes</u>.

This feature of Applicants' invention is of significant utility because, as discussed in the present application, it improves the candidate prefixes and suffixes through iterative refinement.

The other references of record have been reviewed, and these other references, whether considered individually or in combination, also do not disclose or suggest the refinement process of the present invention.
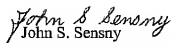
Independent Claims 1, 13 and 21 are being amended to describe this feature of the present invention. Specifically, each of Claims 1, 13 and 21, as presented herewith, describes the feature that the sets of candidate prefixes and suffixes are refined, including using knowledge of suffixes previously identified in the refining to further refine the set of candidate prefixes, and using knowledge of prefixes previously identified in the refining to further refines the set of candidate suffixes.

Because of the above-discussed differences between Claims 1, 13 and 21 and the prior art, and because of the advantages associated with those differences, Claims 1, 13 and 21 patentably distinguish over the prior art and are allowable. Claims 2-12 are dependent from

Claim 1 and are allowable therewith. Also, Claims 14, 15, 19, 20 and 29 are dependent from Claim 13 and are allowable therewith; and Claims 22, 23, 27 and 28 are dependent from, and are allowable with, Claim 21.

The Examiner is, accordingly, respectfully asked to reconsider and to withdraw the rejection of Claims 1-15, 19-23 and 27-29 under 35 U.S.C. 103, and to allow these claims. If the Examiner believes that a telephone conference with Applicants' Attorneys would be advantageous to the disposition of this case, the Examiner is asked to telephone the undersigned.

Respectfully Submitted,

John S. Sensny
John S. Sensny
Registration No. 28,757
Attorney for Applicant

Scully, Scott, Murphy & Presser, P.C.
400 Garden City Plaza, Suite 300
Garden City, New York 11530
(516) 742-4343

JSS:gc

I:\IBM\105\16751\amend\16751.am3.doc